AMINO ACIDS

Amino acids are composed of nitrogen, carbon, hydrogen and oxygen; they comprise a variable side chain group. Our body needs 20 proteinogenic amino acids to grow and function properly. **Proteinogenic** amino acids mean they are required in synthesis of proteins of our body. Only these 20 amino acids are coded by DNA, the genetic material of the cell. The rest amino acids, such as β-alanine, ornithine, γ-aminobutyric acid, citrulline, dioxyphenylalanine are **non-proteogenic.**

Amino Acid	Three letter code	One letter code	MW
Alanine	Ala	A	89.09
Arginine	Arg	R	174.20
Asparagine	Asn	N	132.12
Aspartic Acid	Asp	D	133.10
Cysteine	Cys	C	121.16
Glutamic Acid	Glu	E	147.13
Glutamine	Gln	Q	146.15
Glycine	Gly	G	75.07
Histidine	His	H	155.16
Isoleucine	Ile	I	131.18
Leucine	Leu	L	131.18
Lysine	Lys	K	146.19
Methionine	Met	M	149.21
Phenylalanine	Phe	F	165.19
Proline	Pro	P	115.13
Serine	Ser	S	105.09
Threonine	Thr	T	119.12
Tryptophan	Trp	W	204.23

Tyrosine	Tyr	Y	181.19
Valine	Val	V	117.15

I. Classification of amino acids based on charge

Each amino acid has a carboxyl group and a primary amino group. Exception is proline, because it is an **imino acid.** The variable side chain group ("R group") of amino acids is bond to amino acid's α-carbon atom. At a physiological pH close to 7.4, the carboxyl group is negatively charged (COO^-), and the amino group is protonated ($NH3^+$). Since most amino acids have one carboxyl and one amino group, eventually at physiological pH, hydrophobic and polar amino acids are **zero charged.** All amino acids with uncharged polar residues have zero net charge at pH = 7. Amino acids with basic residues (lysine and arginine) accept protons. So, at physiologic pH, the R groups of Arg & Lys are positively charged. In contrast, histidine is weakly basic and uncharged at physiologic pH. However, when histidine is incorporated into a protein, its R group can be either positively charged or neutral, depending on the ionic environment of the protein. This is an important property of histidine, that provides its buffering role in hemoglobin.

Titration curve of alanine - the *sequential dissociation* of protons from the carboxyl and amino groups of alanine:

Each of amino acid groups has a **pK**, that is equal to the pH at which one **half of the protons** have been **removed** from that group.

The **pK_1 is for acidic group** –COOH, whereas the **pK_2 -** for the **amino group** – NH_3^+. The pK_1 of amino acids is approximately **2**, whereas pK_2 of the α-amino group is approximately **9**.

Isoelectric point: At neutral pH, alanine exists predominantly as the dipolar form, in which the amino- and carboxyl groups give zero net charge. The isoelectric point (pI) is the pH, at which an amino acid is electrically neutral, that is, in which the sum of the positive charges equals the sum of the negative

charges. For an amino acid, such as alanine, the **pI is the average** of pK**1** and pK**2** (pI = [2.3 + 9.1]/2 = **5.7).** The pI is, thus, midway between pK1 and pK_2. The pK_a values and the isoelectronic point (pI) are given below for the 20 α-amino acids. pKa_1 is for α-carboxyl group, pK_{a2} – for α-amino group, and pK_{a3} - for side chain group.

Amino acid	**pKa_1**	**pKa_2**	**pKa_3**	**pI**
Glycine	2.34	9.60	---	5.97
Alanine	2.34	9.69	---	6.00
Valine	2.32	9.62	---	5.96
Leucine	2.36	9.60	---	5.98
Isoleucine	2.36	9.60	---	6.02
Methionine	2.28	9.21	---	5.74
Proline	1.99	10.60	---	6.30
Phenylalanine	1.83	9.13	---	5.48
Tryptophan	2.83	9.39	---	5.89
Asparagine	2.02	8.80	---	5.41
Glutamine	2.17	9.13	---	5.65
Serine	2.21	9.15	---	5.68
Threonine	2.09	9.10	---	5.60
Tyrosine	2.20	9.11	---	5.66
Cysteine	1.96	8.18	---	5.07
Aspartic acid	1.88	9.60	3.65	2.77
Glutamic acid	2.19	9.67	4.25	3.22
Lysine	2.18	8.95	10.53	9.74
Arginine	2.17	9.04	12.48	10.76
Histidine	1.82	9.17	6.00	7.59

In proteins, almost all of these carboxyl and amino groups are joined forming peptide bond and, in general, are not used in chemical reactions. They take part only in hydrogen bond formation (in secondary structure). Thus, it is the nature of the side chains that ultimately dictates the role, which an amino acid plays in a protein. It is, therefore, useful to classify the amino acids ***according to the properties of their side chain:***

1. Hydrophobic amino acids, in which R-groups are non-polar:

Glycine, Alanine, Valine, leucine, Isoleucine, Methionine, Phenylalanine, Proline, Tryptophan

2. Hydrophilic, uncharged amino acids, in which R-group is polar, hydrophilic but uncharged:

Cystein, Serine, Threonine, Tyrosine, Glutamine, Asparagine

3. Acidic amino acids, in which R-group is acidic, negatively charged:

Glutamic acid and Aspartic acid

4. Basic amino acids - amino acids, in which R-group is basic, positively charged:

Lysine, Arginine,

5. Histidine has a variable polarity. Depending on the environment, it can be negatively, positively charged, or remain neutral.

II. Classification of amino acids on the basis of nutrition value.

All 20 of them are important for our health, but among them, nine are absolutely indispensable (essential), since they cannot be produced by the human body and must come from food. The best sources of essential amino acids are animal proteins like meat, eggs and poultry.

When we eat the protein, it is digested to amino acids, which are then used to help the body with such processes as building muscle, regulating immune function etc.

1. Essential amino acids:

The essential amino acids are not synthesized in humans, but they are synthesized in plants or bacteria. Plants are able to make all the amino acids. But humans do not have all the the enzymes required for the biosynthesis of these amino acids: they are required in the diet. These 9 essential amino acids are histidine (His), isoleucine (Ile), leucine (Leu), methionine (Met), phenylalanine (Phe), threonine (Thr), tryptophan (Trp), and valine (Val):

Leucine, Isoleucine, Phenylalanine, Histidine, Valine, Threonine, Tryptophan, Methionine, Lysine

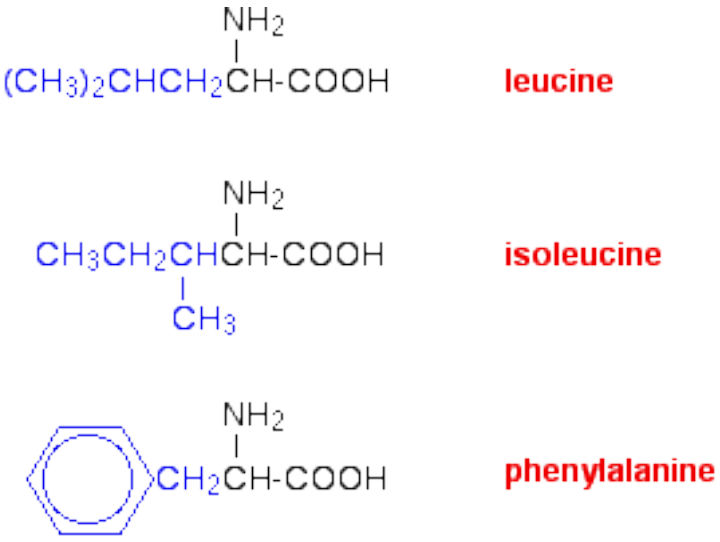

Fig. 1. Some hydrophobic essential amino acids

Since these amino acids are not synthesized in cells of human, so they should be present in the diet. Although histidine is synthesized by the gut microbiota, its supply does not meet our needs, so for a healthy lifestyle, histidine must be supplemented additionally with food.

The essential amino acids are involved in many pathways and processes that are crucial for the homeostasis of the body.

In conclusion, the essential amino acids are needed to produce other non-essential amino acids. They are also involved in a gene regulation and ATP generation.

2. Conditionally essential amino acids

There are several non-essential amino acids that are classified as conditionally essential, since they are essential only under specific circumstances such as illness or stress. Arginine is conditional amino acids. It is essential for children, non essential for adults. Although arginine is non-essential, organism can't meet demands when fighting certain diseases like cancer: arginine must be supplemented through diet in order to meet the body's needs in certain situations.

3. Non-essential amino acids.

The rest 10 amino acids, that we can produce are:
alanine, asparagine, aspartic acid, cysteine, glutamic acid, glutamine, glycine, proline, serine and tyrosine.

Tyrosine is produced from phenylalanine, so if the diet is deficient in phenylalanine, tyrosine will be required.

These aminoacids can be synthesized in body, so they need not be included in diet.

Glycine, Alanine, Serine, Cysteine, Asparagine, Glutamine, Aspartic acid, Glutamic acid, Tyrosine, Proline

Functions of some amino acids:

Methionine is a *methyl group donor* in the synthesis of choline, adrenaline.

Therefore, it has a lipotropic effect in atherosclerosis and fatty infiltration of the liver.

Glycine and β-alanine are inhibitory mediators of the brain. β-alanine and α-alanine are formed by the decarboxylation of aspartic acid. **β-alanine** is constituent part of anserine and carnosine, HS-CoA.

Aspartic acid is involved in the synthesis of purines and pyrimidines. Amino group of glutamin also is used for the synthesis of purines and pyrimidines.

Glutamic acid is part of glutathione (along with cysteine and glycine). Glutamic acid is involved in the synthesis of porphyrins and GABA –γ-amino butiric acid, an inhibitory mediator of the brain.

Lysine and arginine are prevalent in protamines and histones. The ε-amino group of **lysine** binds biotin to its enzyme.

Tyrosine can be converted to thyroxin, adrenaline and noradrenaline, the pigment melanin.

From **tryptophan** with the participation of vitamin $B_{6,}$ nicotinic acid is produced. Therefore, the absence of tryptophan, PP and B_6 vitamins is expressed by pellagra disease.

Histidine in hemoglobin binds protein with iron.

Elastin and collagen are rich in **prolin**.

Structural levels of proteins.

PRIMARY STRUCTURE OF PROTEINS

Since each amino acid has a carboxyl group and an amine group, amino acids link to one another to form a chain via peptide bond, which joins the amine group of one amino acid to the carboxyl group of the next. Proteins are naturally synthesized starting from the N-terminus and ending at the C-terminus. Thus, polypeptide chain has an end of initial amino acid with

an unbound amine group, the N-terminus, and a last amino acid with an unbound carboxyl group, which is termed the C-terminus.

A peptide bond is a covalent bond, that is formed between two neighboring amino acids. This chain composed from amino acids is termed polypeptide chain, or protein. A peptide is a chain of amino acids in which the **α-amino group** of one amino acid is bonded to the **α-carboxyl group** of the next. Each bond linking two neighboring amino acids is called a **peptide bond**.

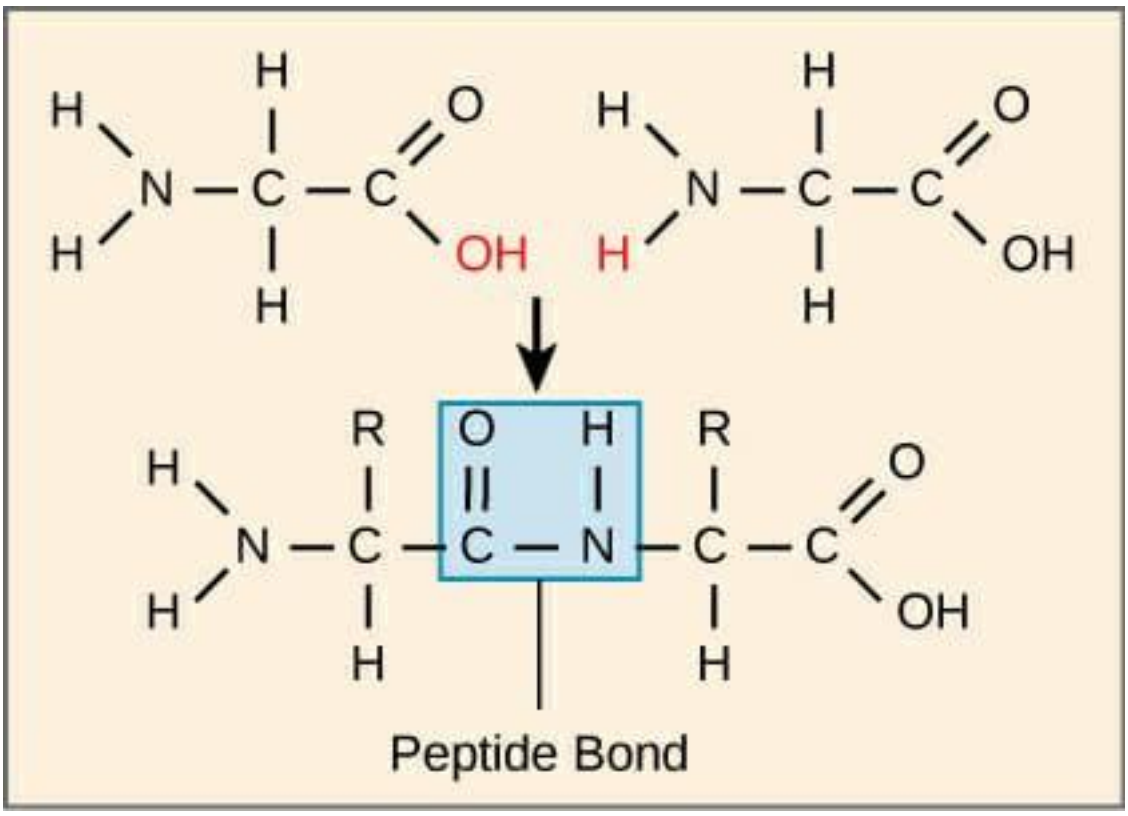

Fig.2. Peptide bond formation

If a peptide is made from two amino acids, it is termed a dipeptide; made from three amino acids is called a tripeptide, and so on. The prefixes, *di-*, *tri-*, *tetra-* indicate the number of amino acid units in peptide. Peptides that contain up to fifty amino acids are called peptides. Peptides with more than 50 amino acids are called proteins.

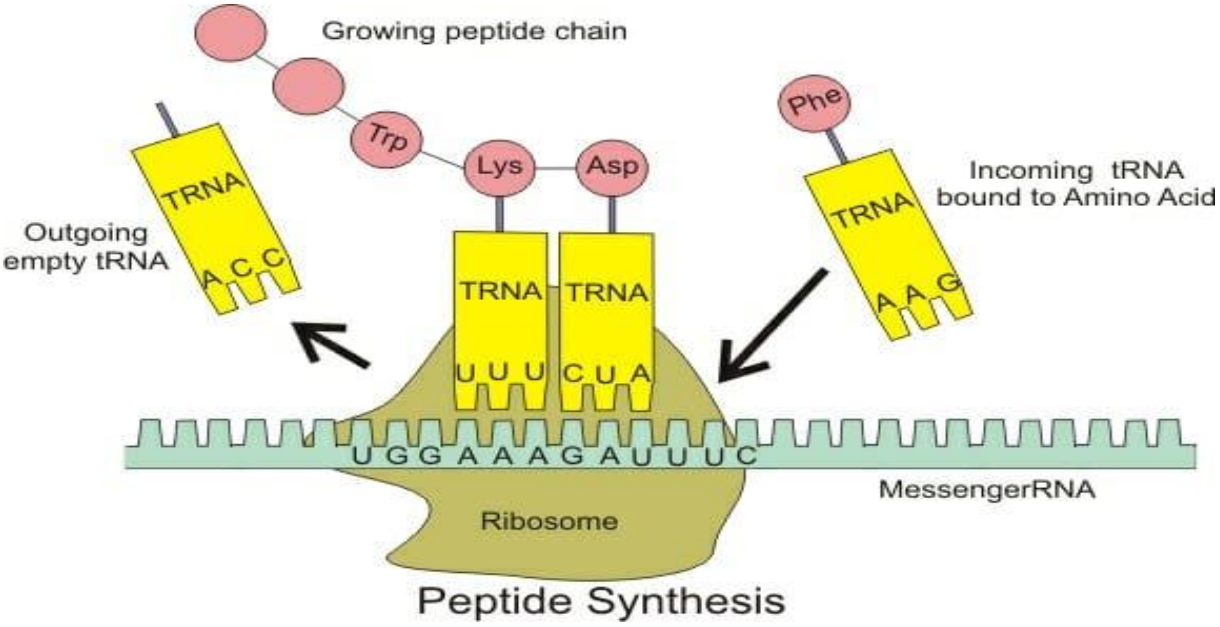

Fig.3. Sequential incorporation of amino acids into polypeptide chain on ribosomes.First amino acid is N-teminal.

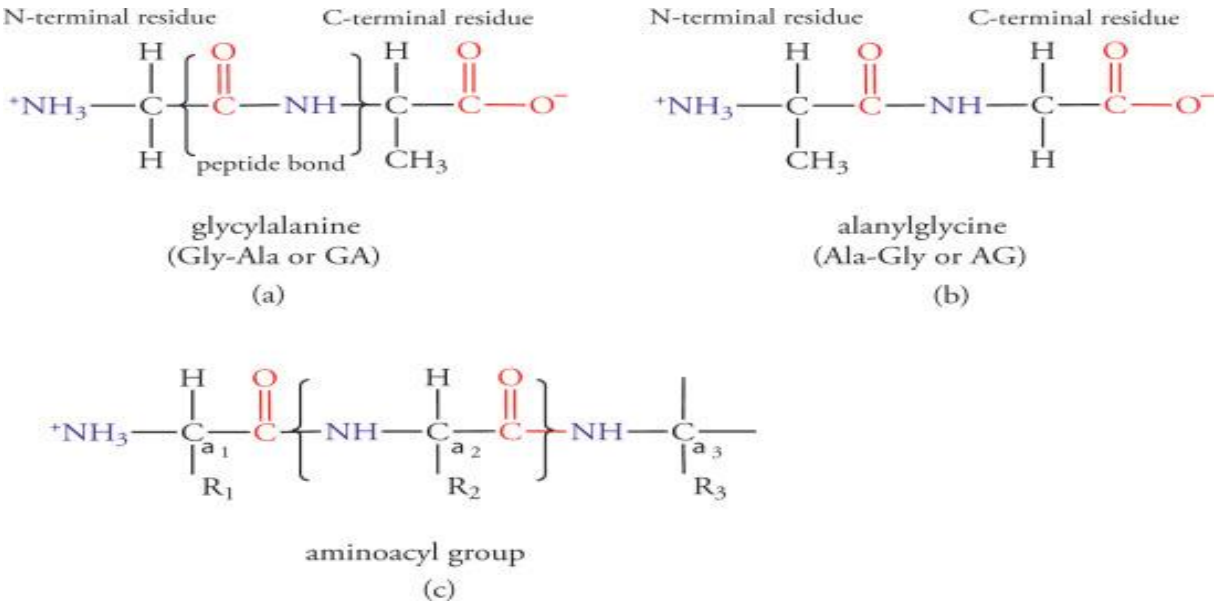

Fig. 4. The scheme shows the N-terminal & C-terminal amino acids and the name of the peptides

The unique geometry of proline contributes to the formation of the fibrous structure of collagen.

Proteins have four organizational levels: primary, secondary, tertiary and quaternary. The primary structure of a protein is the sequence of amino acids in a protein. Amino acids are bound in the primary structure of a protein

through peptide bond. The biuret reaction is characteristic for a peptide bond. All proteins and amino acids give positive biuret and ninhydrin reactions.

When a polypeptide is named, all amino acids change their suffixes (-ine, -an, -ic, or -ate) to **-yl**, with the exception of the last, C-terminal amino acid. For example, a tripeptide composed of an N-terminal glycine, leycine and a C-terminal alanine is called glyc**yl**-leuc**yl**-alanine.

Sickle cell anemia, a sickling disease of red blood cells, results from the replacement of polar glutamate with nonpolar valine at the sixth position in the β- subunit of hemoglobin.

Determination of the amino acid composition of a polypeptide.

Sequencing is a stepwise process of identifying the specific amino acid at each position in the peptide chain, beginning at the N-terminal end. N-terminal amino acid is determined by the method of Sanger, Edman, dansyl chloride, aminopeptidase.

Aminopeptidases are used in determination of amino acids sequence in peptide chain, since they catalyze the hydrolysis of peptide removing N-terminal amino acid from N-trminus.

C-terminal amino acid is determined with carboxypeptidases and Akabori method. The **C** is the end of an amino acid chain terminated by a free carboxyl group. When the protein is translated from messenger RNA, it is created from N-terminus to C-terminus. The convention for writing peptide sequences is to put the C-terminal end on the right and write the sequence from N- to C-terminus.

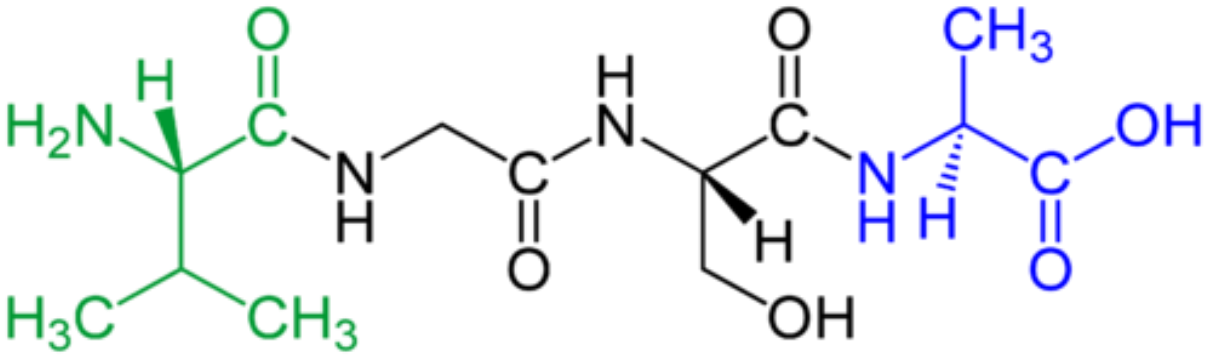

Fig. 5. A tetrapeptide example: Val-Gly-Ser-Ala. Green is highlighted *N*-terminal amino acid (valine), blue is marked *C*-terminal α-amino acid (alanine).

While the N-terminus of a protein often contains targeting signals, the C-terminus can contain retention signals for protein sorting. The most common ER retention signal is the amino acid sequence KDEL (Lys-Asp-Glu-Leu) or HDEL (His-Asp-Glu-Leu) at the C-terminus. This keeps the protein in the endoplasmic reticulum and prevents it from entering the secretion.

SECONDARY STRUCTURE OF PROTEINS

In the secondary structure, the polypeptide backbone folds into the **α**-helix or **β**-sheet.

An *α-helix* is stabilized by **hydrogen bonds** between *the peptide-bond carbonyl oxygens* and *amide hydrogens* that are part of the polypeptide backbone. The *hydrogen bonds are located parallel* to the spiral helix. So all, but the first and last peptide bond components are linked to each other through intrachain hydrogen bonds. Hydrogen bonds are individually weak, but they collectively serve to stabilize the helix. Each turn of an α-helix contains 3.6 amino acids. ***Proline*** disrupts an α-helix because its secondary amino group is not geometrically compatible with the right-handed spiral of the α-helix. Instead, it ***inserts a kink*** in the chain.

Structure of a β-sheet.

The ***hydrogen bonds are perpendicular*** to the polypeptide backbone in β-sheets. Because peptide chains have a directionality conferred by their N-terminus and C-terminus, β-strands too can be said to be directional. They are usually represented in protein topology diagrams by an arrow pointing toward the C-terminus. Adjacent β-strands can form hydrogen bonds in antiparallel, parallel, or mixed arrangements.

In an antiparallel arrangement, the successive β-strands alternate directions so that the N-terminus of one strand is adjacent to the C-terminus of the next. This is the arrangement that produces the strongest inter-strand stability because it allows the inter-strand hydrogen bonds between carbonyls and amines to be planar, which is their preferred orientation.

Antiparallel β-sheet hydrogen bonding patterns, represented by dotted lines. Oxygen atoms are colored **red** and nitrogen atoms colored **blue**

In an antiparallel arrangement, the successive β-strands alternate directions so that the N-terminus of one strand is adjacent to the C-terminus of the next. This is the arrangement that produces the strongest inter-strand stability because it allows the inter-strand hydrogen bonds between carbonyls and amines to be planar, which is their preferred orientation

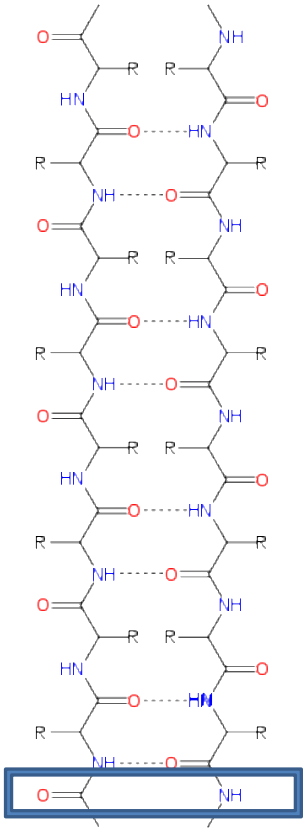

Fig. 6. Scheme of a antiparallel β-sheet structure.

In a parallel β-sheet arrangement, all of the N-termini of successive strands are

oriented in the same direction; this orientation may be slightly less stable because it introduces nonplanarity in the inter-strand hydrogen bonding pattern.

Parallel β-sheet hydrogen bonding patterns, represented by dotted lines. Oxygen atoms are colored **red** and nitrogen atoms colored **blue**

In a parallel arrangement, all of the N-termini of successive strands are oriented in the same direction; this orientation may be slightly less stable because it introduces nonplanarity in the inter-strand hydrogen bonding pattern.

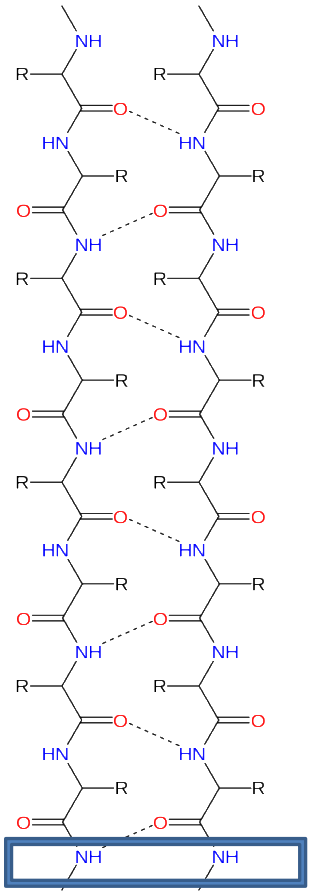

Fig. 7. Parallel β-sheet structure.

Finally, an individual strand may exhibit a mixed bonding pattern, with a parallel strand on one side and an antiparallel strand on the other.

The hydrogen bonding of β-strands need not be perfect, but can exhibit localized disruptions known as β-bulges. Approximately one half of an average globular protein is organized by α-helix and β-sheet. The remainder of the polypeptide chain has a loop or coil conformation, which have a less regular structure than those described above.

The TERTIARY STRUCTURE of PROTEINS.

In the tertiary structure, the protein including secondary structure folds into final three-dimentional shape. It is often simplified in models.

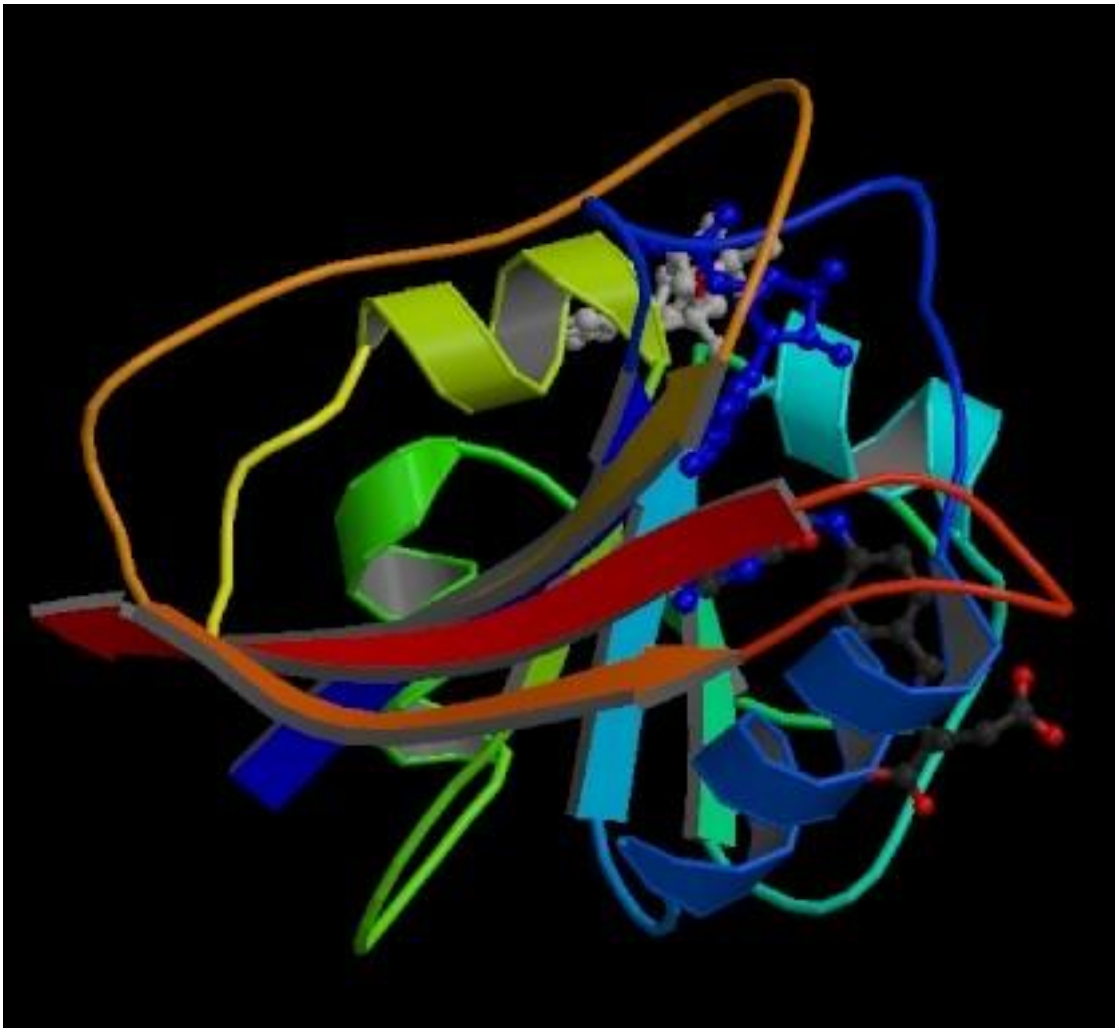

Fig.8. The arrangement of the secondary structures into this final 3-D structure. Dihydrofolate reductase tertiary structure.

The model above shows the alpha-helixes of secondary structure and beta-pleated sheets, Random coils and loops are also present. These patterns are folded vi interactions formed between R-groups of amino acids.

All proteins in tertiary structure have globular or fibrous shape. So, there are 2 types of tertiary structure - 1) globular 2) fibrous

Fibrous proteins are collagen, elastin, keratin, fibroin etc.Their structure in aqueous solution is compact like a globule. Hydrophobic side chains are buried in the interior, whereas hydrophilic groups are generally found on the surface of the globular proteins molecule.

Most of proteins are globular, only some of them, particularly scleroproteins

have fibrous form.

Most proteins fold into their tertiary structure in an aqueous environment - a cell is, after all, 60% water. The chemical properties of the various R-groups (side chains) of the amino acids within the protein chain will influence the way that the protein folds in its environment.When a protein is surrounded by water:

- Hydrophobic amino acids will move away from the water and bury themselves in the center of the protein.
- Hydrophilic amino acids will interact with the water molecules, and thus tend to be located on the outer surface of the protein.
- Basic (positvely charged) amino acids and acidic (negatively charged) amino acids create salt bridges, or electrostatic interactions, which further stabilize the tertiary structure.
- Cysteines may form a disulfide bridge, further stabilizing the protein.

Bonds stabilizing tertiry structure

Interactions between the amino acid side chains guide the folding of the polypeptide to form a compact structure. The following four types of interactions cooperate in stabilizing the tertiary structure of globular proteins:

1. *Disulfide bonds*: A disulfide bond is a covalent linkage formed from the sulfhydryl group (–SH) of of two cysteine residues to produce a cystine residue. So, disulfide bond is formed between cysteine residues and participates in the formation of all - primary, secondary, tertiary, quaternary protein structures.
2. *Hydrophobic interactions*: Amino acids with nonpolar side

chains tend to be located in the interior part of the polypeptide molecule, where they associate with other hydrophobic amino acids. In contrast, amino acids with polar or charged side chains tend to be located on the surface of the molecule in contact with the polar solvent.

3. *Hydrogen bonds:* Amino acid side chains containing oxygen- or nitrogen-bound hydrogen, such as serine and threonine, can form hydrogen bonds with oxygen of a carboxyl group or carbonyl group. Lots of **amino acids** contain groups in the side chains which have a **hydrogen** atom attached to either an oxygen or a nitrogen atom. This is a classic situation where **hydrogen bonding** can occur. Serine, threonine, and tyrosine each contain a polar **hydroxyl group**, that can participate in **hydrogen bond** formation. You could have a **hydrogen bond** set up between two serine residues in different parts of a folded chain. The residues of **asparagine and glutamine** can also participate in hydrogen bonds).

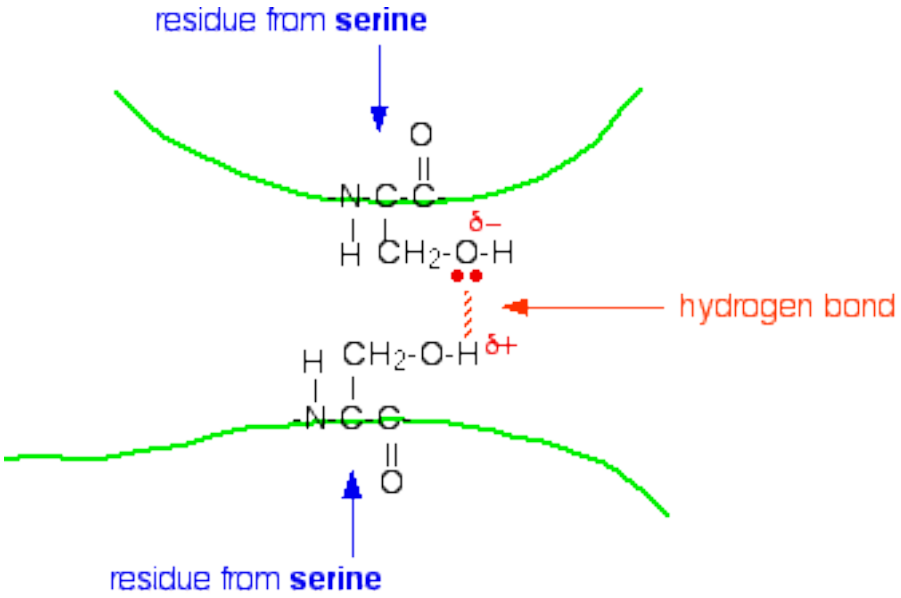

Fig. 10. Hydrogen bond between two serine OH-groups

4. *Ionic interactions:* Negatively charged groups, such as the carboxylic group COO– of aspartate or glutamate, can interact with positively charged amino group NH_3^+ of lysine.

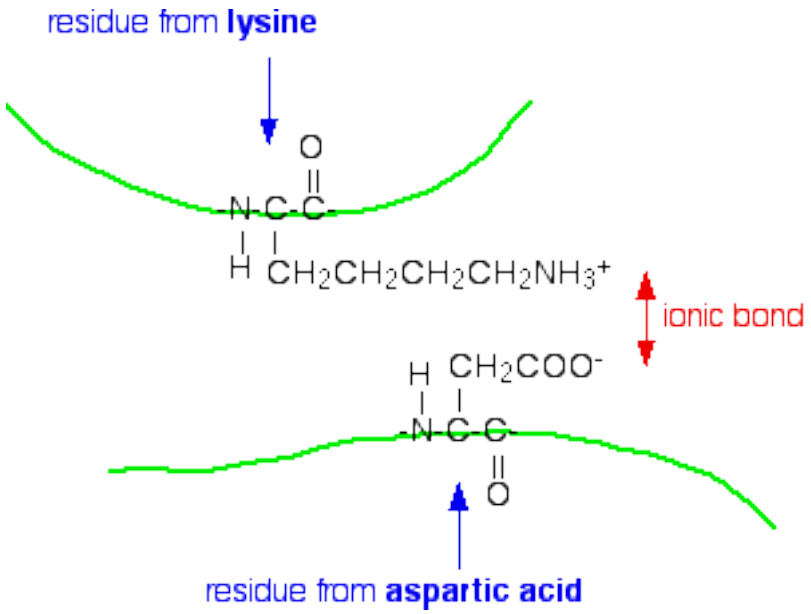

Fig. 10. Ionic bond between negatively and positively charged amino acids

Protein folding

Active proteins are used in many roles including structural support, catalyzing enzymatic reactions, and recognizing molecules.Interactions between the residues of amino acids determine how a long polypeptide chain folds into the intricate three-dimensional shape of the functional active protein. Protein folding, which occurs within the cell in seconds to minutes, involves non random, but ordered pathways. As a peptide folds, secondary structures

form.These small structures combine to form larger structures. Additional events stabilize secondary structure and initiate formation of tertiary structure. In the last stage, the peptide achieves its fully folded, native, functional active form characterized by a low energy state. Some biologically active proteins or segments lack a stable tertiary structure. They are referred to as "intrinsically disordered" proteins.

Role of chaperones in protein folding.

The information needed for correct protein folding is contained in the primary structure of the polypeptide. However, most proteins when denatured do not resume their native conformations even under favorable environmental conditions. This is because, for many proteins, folding is a facilitated process that requires a specialized group of proteins, referred to as "molecular chaperones," and adenosine triphosphate hydrolysis. The chaperones, also known as "heat shock proteins" (HSP), interact with apopeptide at various stages during the folding process. Some chaperones bind hydrophobic regions of an extended polypeptide and are important in keeping the protein unfolded until its synthesis is completed. Others form cage-like macromolecular structures composed of two stacked rings. The partially folded protein enters the cage, binds the central cavity through hydrophobic interactions, folds, and is released (for example, mitochondrial HSP 60). Cage-like chaperones are also termed "chaperonins." Chaperones, then, facilitate correct protein folding by binding to and stabilizing exposed, aggregation-prone hydrophobic regions in nascent and denatured polypeptides, preventing premature folding.

Protein folding sometimes result in improperly folded molecules. These misfolded proteins are usually tagged and degraded within the cell . However, this quality control system is not perfect, and intracellular or extracellular aggregates of misfolded proteins can accumulate, particularly as individuals age.

Deposits of misfolded proteins are associated with a number of diseases.

Amyloid disease

Misfolding of proteins may occur spontaneously or be caused by a mutation in aparticular gene, which then produces an altered protein. In addition, some apparently normal proteins can, after abnormal proteolytic cleavage, take on a unique conformational state that leads to the formation of long, fibrillar protein assemblies consisting of β-pleated sheets. Accumulation of these insoluble, spontaneously aggregating proteins, called amyloids, has been implicated in degenerative diseases such as Parkinson and Huntington and particularly in the age-related neurodegenerative disorder, Alzheimer disease.

QUATERNARY STRUCTURE OF PROTEINS.

Many proteins consist of a single polypeptide chain, however, others may consist of two or more polypeptide chains that may be structurally identical or totally unrelated. The arrangement of these polypeptide subunits is called the quaternary structure of the protein. Subunits are held together primarily by noncovalent interactions, for example, hydrogen bonds, ionic bonds, and hydrophobic interactions. Subunits may either function independently of each other or may work cooperatively, as in hemoglobin, in which the binding of oxygen to one subunit increases the affinity of the other subunits for oxygen.

Denaturation of proteins.

During denaturation all the stucture levels (secondary, tertiary, and quaternary) with exception of primary structure are destroyed. The primary structure is destroyed only at ***hydrolysis.*** Denaturation can occur from physical factors, such as heating, radiation, pressure. The chemical agents, such as acids, alkalis, heavy metals, organic solvents such as alcohol, acetone, detergents can

aso denature the proteins. Urea and guanidinum chloride break hydrogen bonds. Acids and alkalis break down ionic bonds. At 3> pH> 10, most proteins denature. But pepsin is denatured at pH = 6.

Conditions that denature proteins:

1. **Extreme pH** (3> pH > 10) : alters H-bonding.

2. **Heat** (temp >50^{o}C): thermal effect, disrupts weak forces of non-covalent bonds.

3. **Detergents or organic solvents** : disrupts hydrophobic interaction

4. **Chaotropic agents** (high concentrations) : e.g., urea and guanidinium chloride

Denatured protein produces more intense color reactions. Denaturation may, under ideal conditions, be reversible, such that the protein refolds into its original native structure when the denaturing agent is removed. So, renaturation is the restoration of the physico-chemical and biological properties of denatured protein.However, most proteins, once denatured, remain permanently disordered. Denatured proteins are often insoluble and precipitate from solution.

There are not only proteins, but also natural peptides in the body, which play an important role in the regulation of metabolism and physiologicak functions of body. These peptides are sort into 4 groups:

1) hormonal,
2) digestive,
3) neuropeptides,
4) derivatives of α_2-globulin fraction.

Hormone-peptides are vasopressin, oxytocin, glucagon, calcitonin. Both vasopressin and oxytocin are composed of only 9 amino acids.

Digestive peptides are gastrin, enterogastron, secretin, pancreozymin.

Peptides derived from **α_2-globulin** fraction are angiotensin, bradykinin, kallidin.

renin

angiotensinogen ⟶ angiotensin I

↓ *carboxy-cathepsin*

angiotensin II

Carboxy-cathepsin cleaves 2 amino acids from C-end of angiotensin I converting it to angiotensin II. Angiotensin II is the strongest vasopressor, enhances the synthesis of aldosterone, which increases the reabsorption of Na^+ and water in the kidneys, therefore rises a blood pressure.

Bradykinin is a vasodilator, formed from callidin.There are 10 amino acids in callidin and 9 in bradykinin. Synthesis of these kynines occur following way:

Prekallikrein ⟶ kallikrein

↓

Kininogen → Lysyl-Bradykinin (Kallidin)

↓ *Aminopeptidase*

BRADYKININE

Aminopeptidase cleaves 1 amino acid from kallidin, forming the strongest vasodilator bradykinin. It increases in the blood at inflammation and allergies.Bradykinin dilates blood vessels, reduces pressure, increases vascular permeability. Carboxycatepsin cleaves the dipeptide from bradykinin and thus reduces its activity.

Glutathione, anserine and carnosine are also natural peptides.

Glutathione present in erythrocytes, kidneys, liver, where it is necessary to neutralize H_2O_2.

Anserine is a methylated carnosine. These dipeptides restore the pH in a tired muscle and rise its activity.

Neuropeptides are the hormonal factors of the hypothalamus - ***thyroliberin***, melanoliberin, melanostatin***, enkephalins*** etc.